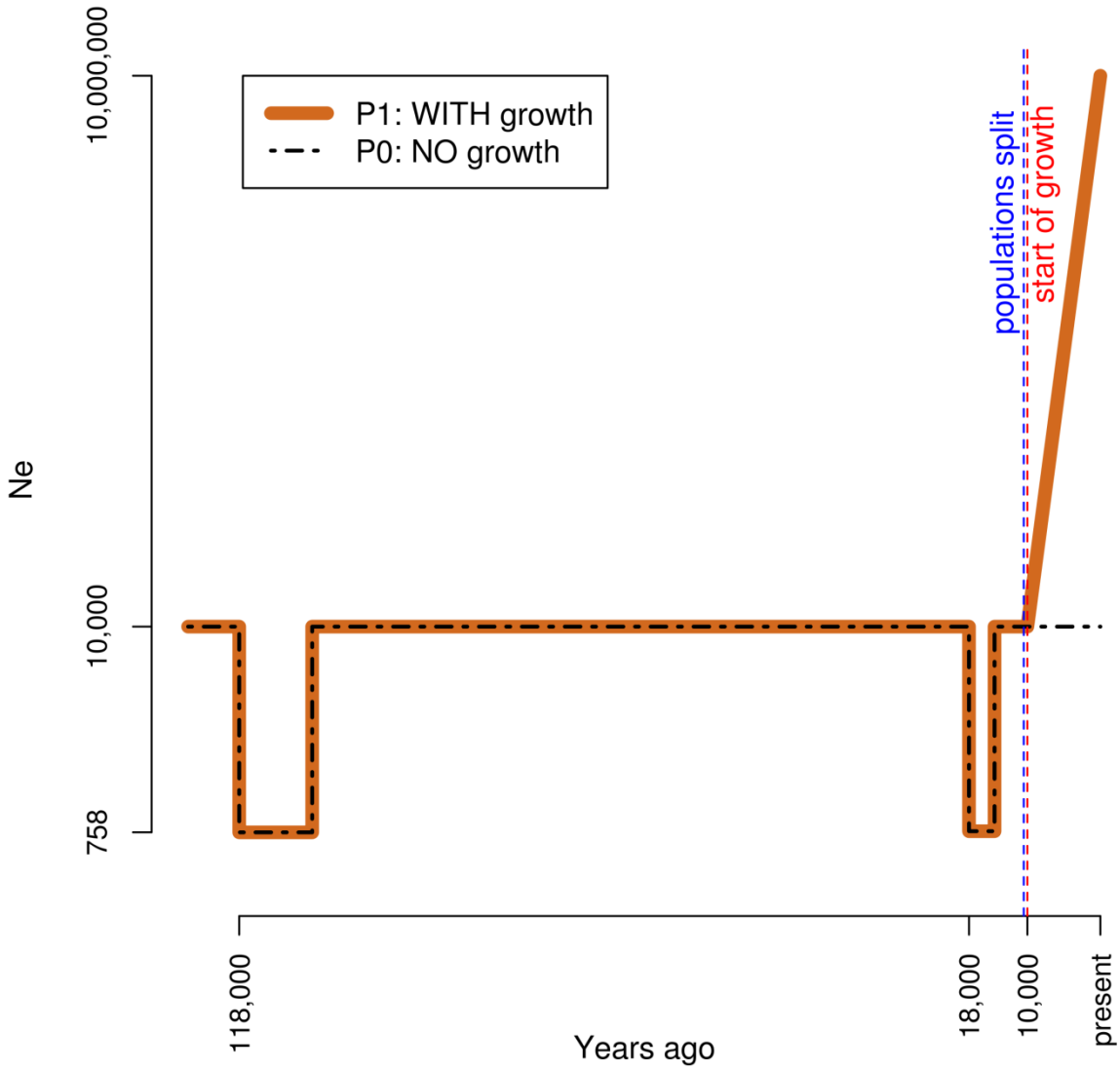


Figure S1 Main demographic models used. The y-axis represents Ne at different epochs (x-axis) on a log scale. Both populations (with growth, no growth) have an ancestral population size of N_e = 10,000. They undergo a first bottleneck of intensity F = 0.264 at 4,720 generations ago followed by a quick recovery to the ancestral population size. The second bottleneck occurs 720 generations ago with an intensity of F = 0.09, also followed by a quick recovery to the ancestral population size. Both populations share a common ancient history until 420 generations ago where a copy (P1) of the ancestral population (P0) is created. Then, 400 generations before present, P1 starts growing at a rate of 1.73% per generation, reaching a final size of N_e = 10,000,000 at the end of the simulation. In contrast, P0 continues to evolve maintaining its size constant at N_e = 10,000 until present.

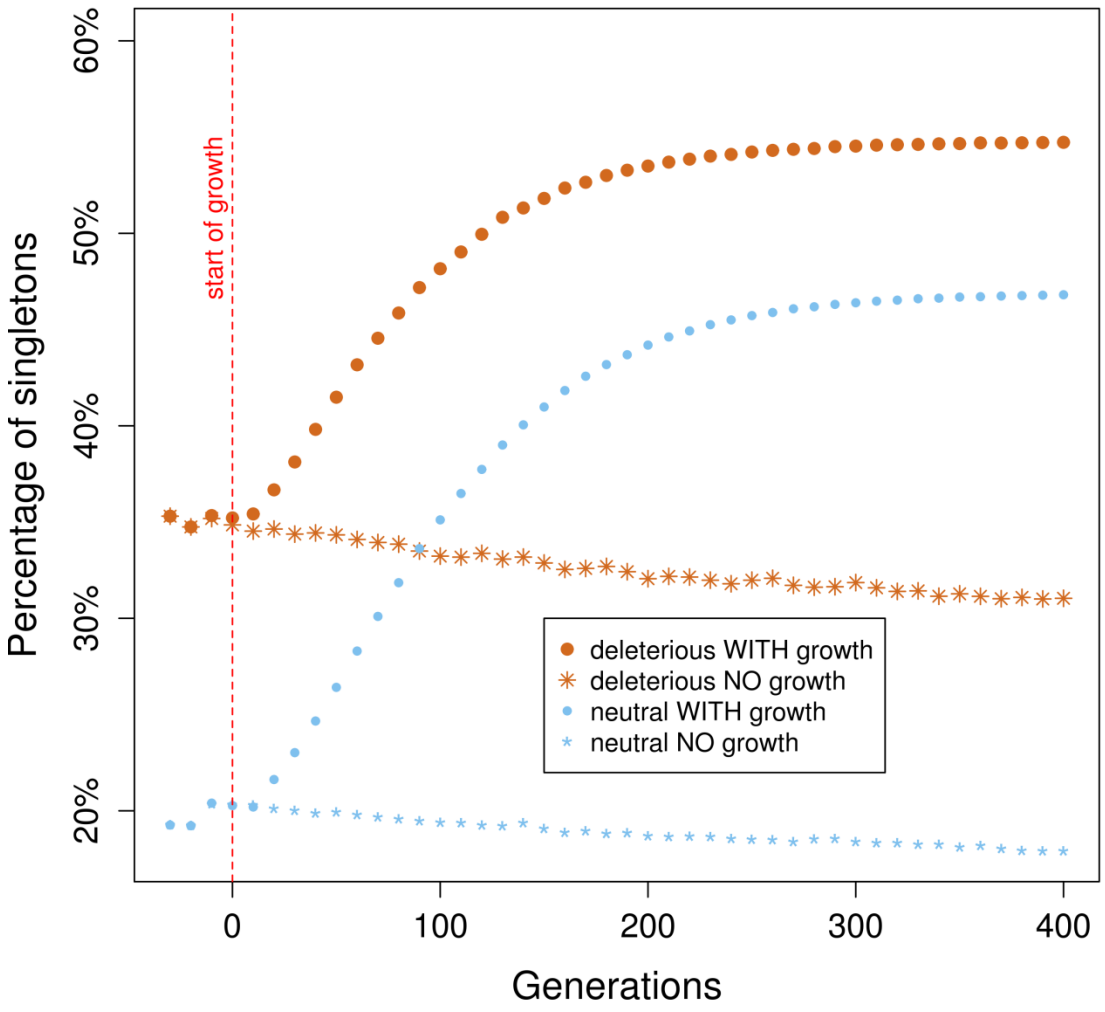


Figure S2 The percentage of singletons increases in a growing population. The percentage of singletons (sites with derived allele count, DAC = 1) out of all the segregating sites is shown for the last 440 generations of the simulation, for loci with either deleterious or neutral mutations, for both population models (NO growth, WITH growth).

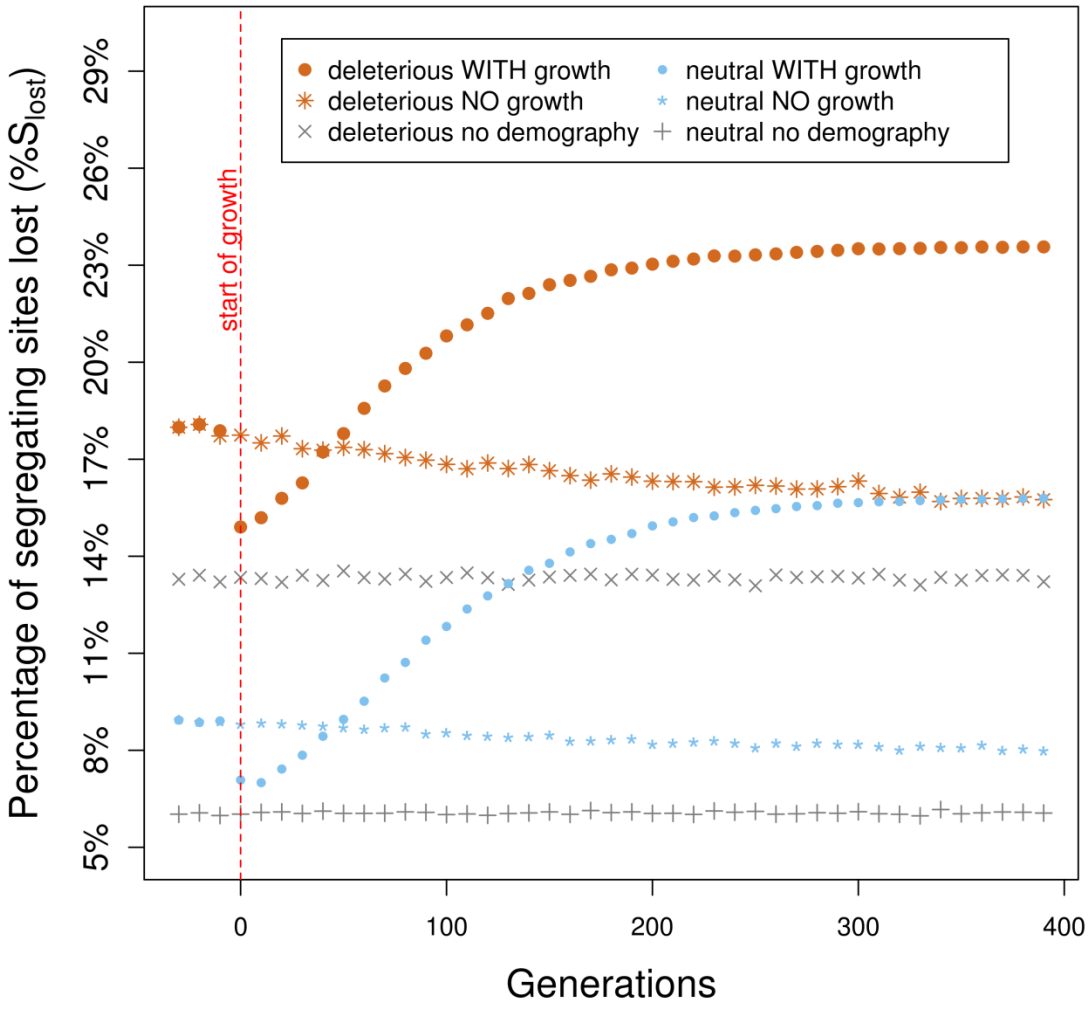


Figure S3 The percentage of segregating sites lost at each generation decreases in population without growth while it is constant in a population without any demography. The growing and constant populations data are the same as presented in Figure 1b. The population with no demography has a constant effective population size of 10,000 throughout history without any demographic events, thus differing from the constant size model by the absence of ancestral bottlenecks. The comparison of the constant size population and the population with no demography (gray) allows assessing the effect of past demographic events (bottlenecks). The percentage of segregating sites lost $\%S_{lost}$ at each generation is slightly decreasing in the population without growth, but not in the scenario without demographic event. This comparison reveals that the population without growth, although evolving at constant size for the last 580 generations has not recovered its mutation-drift balance that was altered during the ancient bottlenecks.

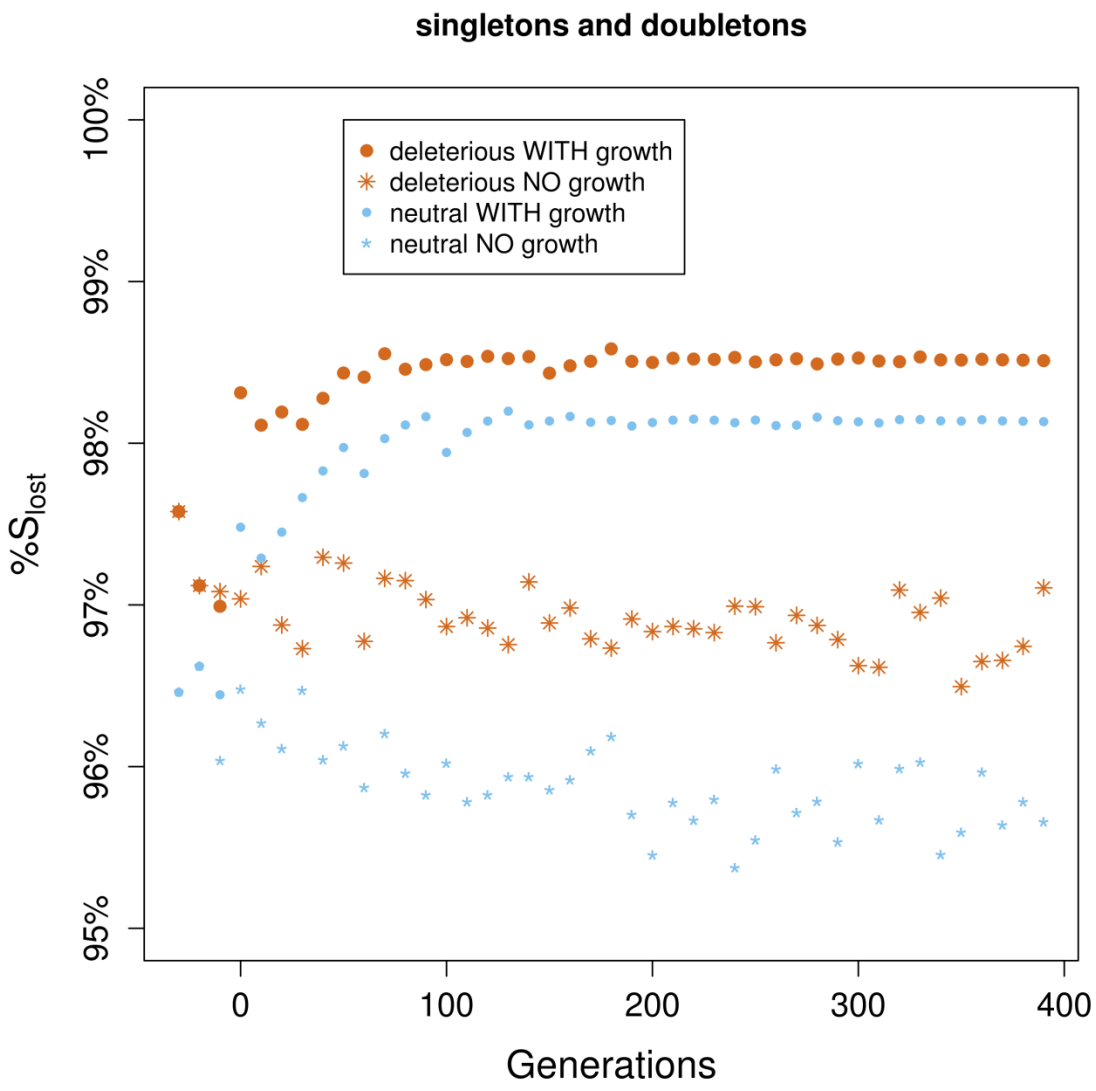


Figure S4 The contribution of singletons and doubletons to lost sites increases in a growing population. Considering all lost segregating sites, we counted the percentage that have a certain count of derived alleles (DAC = 1, DAC = 2, DAC = 3 etc). These percentages for all possible DAC (from 1 to 2 N_e - 1) sum up to 1. In the growing scenario, singletons and doubletons represent over 98% of the derived alleles lost. For higher DAC (>= 3), the contribution of segregating sites with higher DAC decreases quickly as DAC increases. For example, all sites with DAC>10 combined only represent between 0 and 0.00016% of the sites that are lost at a deleterious locus in a growing population, depending on the generation considered (data not shown).

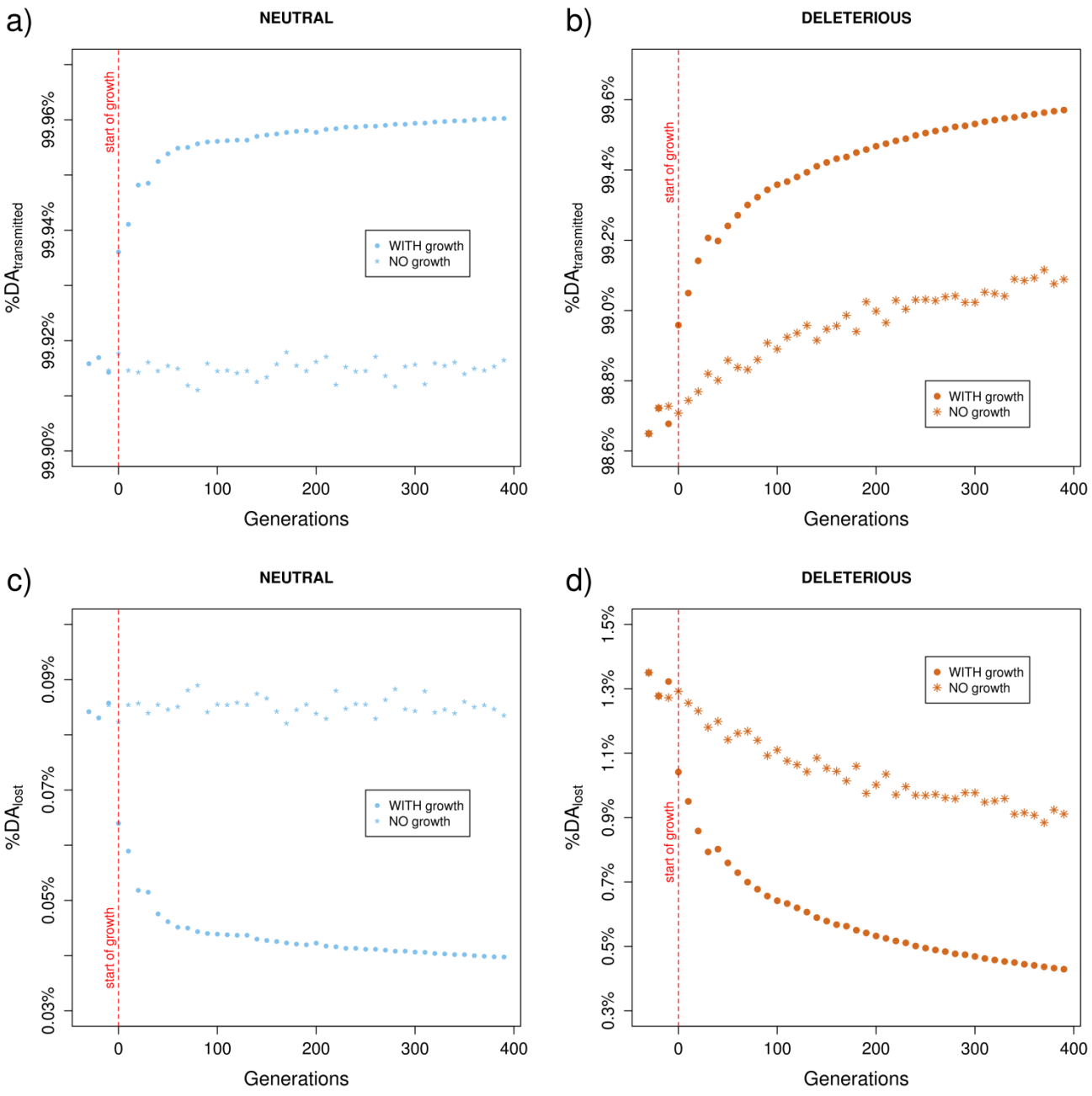


Figure S5 The fraction of derived alleles lost under population expansion is higher at a deleterious locus than at a neutral locus. The fraction of derived alleles lost, $\%DA_{lost}$, is defined as the sum of the number of copies of derived alleles at the segregating sites lost, over the sum of the number of copies of derived alleles present at all segregating sites present in the population. The fraction of derived alleles transmitted, $\%DA_{transmitted}$, is 1-$\%DA_{lost}$. The $\%DA_{transmitted}$ (a, b) and the $\%DA_{lost}$ (c, d) are shown for each locus type, neutral (a, c) or deleterious (b, d). The growing populations lose a higher percentage of derived alleles at the deleterious locus than at the neutral locus.

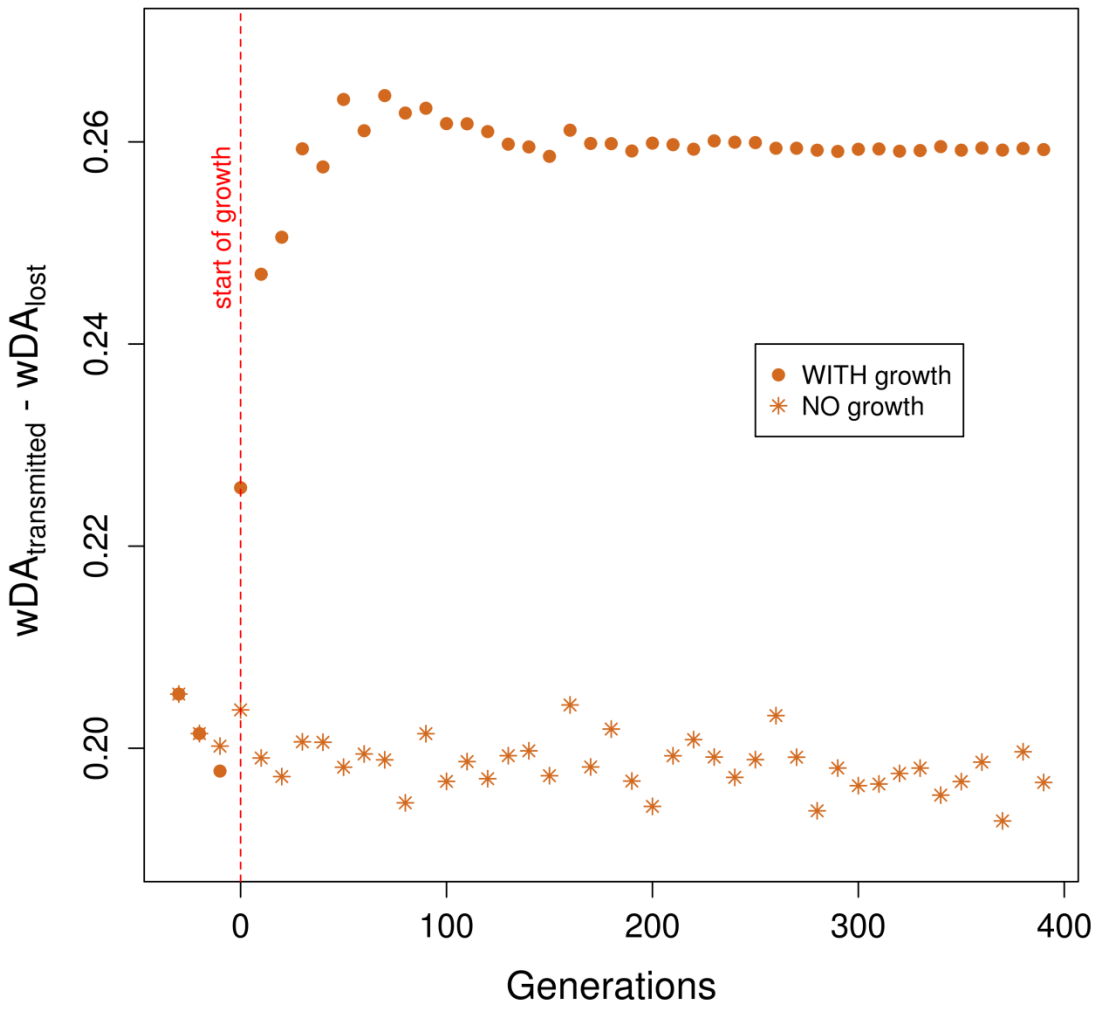


Figure S6 The difference in fitness effect between lost and transmitted derived alleles increases in a growing population. The average fitness effect (*wDA*) is defined as in Figure 4. The difference in fitness effect is the difference between *wDA* at transmitted sites and *wDA* at lost sites ($wDA_{transmitted}$ - wDA_{lost}). All the mutations are deleterious and have a negative fitness effect. Positive values on the y-axis show that the average fitness at transmitted sites is less deleterious than at lost sites. The difference $wDA_{transmitted}$ - wDA_{lost} is higher in the growing population, showing that the fraction of alleles that are transmitted have on average a better fitness effect than the ones that are lost in this demographic scenario compared to a scenario without growth.

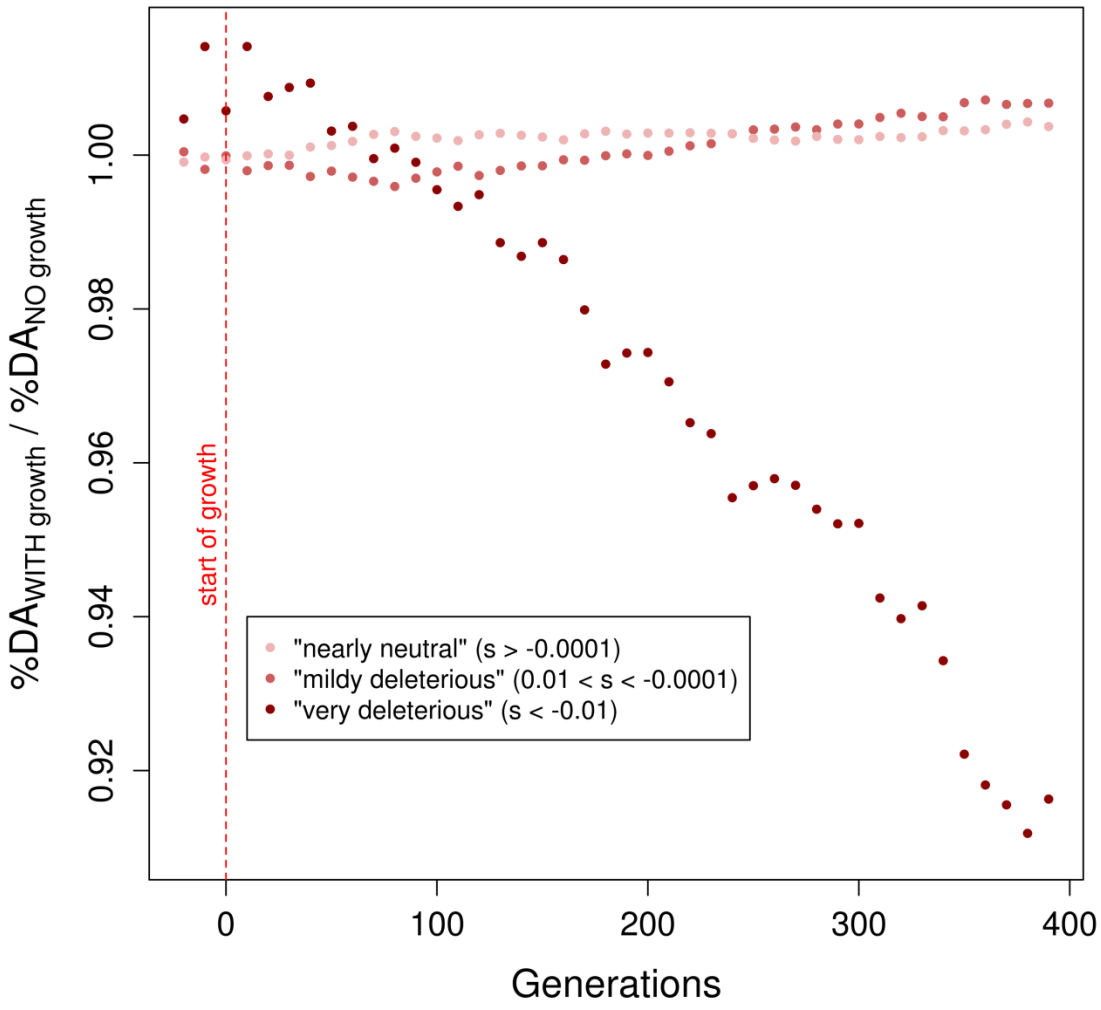


Figure S7 The percentage of derived alleles in the most deleterious category of fitness effect decreases in the growing population. The percentage of derived alleles (*%DA*) and the category of fitness effect are defined as in Figure 3. The data are presented as the ratio of *%DA* in the growing population over the *%DA* of the population of constant size. The ratio below 1 in the most deleterious category indicates that a lower percentage of derived alleles have a fitness effect < -0.01 in the scenario with growth.

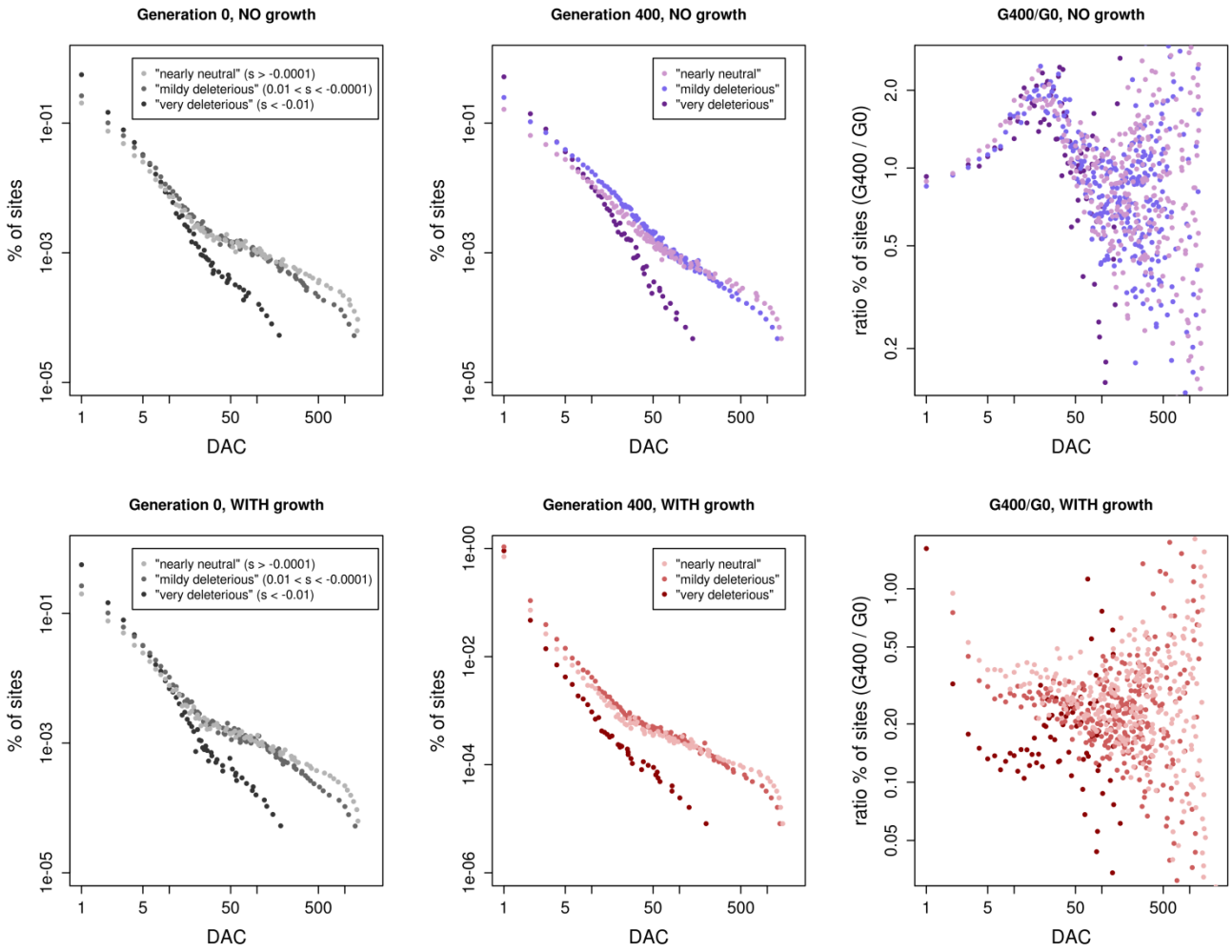


Figure S8 Population growth mainly affects the most deleterious allele category. The site frequency spectrum (SFS) of the population is shown before (a) and after 400 generations growth (b), for three different categories of fitness effect. During the same interval of time, we also show the three same SFS for a population that has not experienced the growth (d,e). Data are presented on a log-log scale. Generation 400 represents present time and generation 0 represents the time at which one population starts growing. Panels (c) and (e) show the ratio of the proportion of sites for each derived allele count (DAC) at time 400 over time 0. For DAC greater than 20, data are noisy because counts are low in each DAC. We observe that population growth has induced a skew of the SFS toward rare variants (ratio below 1). In addition, this skew is more accentuated for the alleles in the most deleterious category. In contrast, the SFS for the three categories of fitness have changed almost identically after 400 generation in the population without growth.

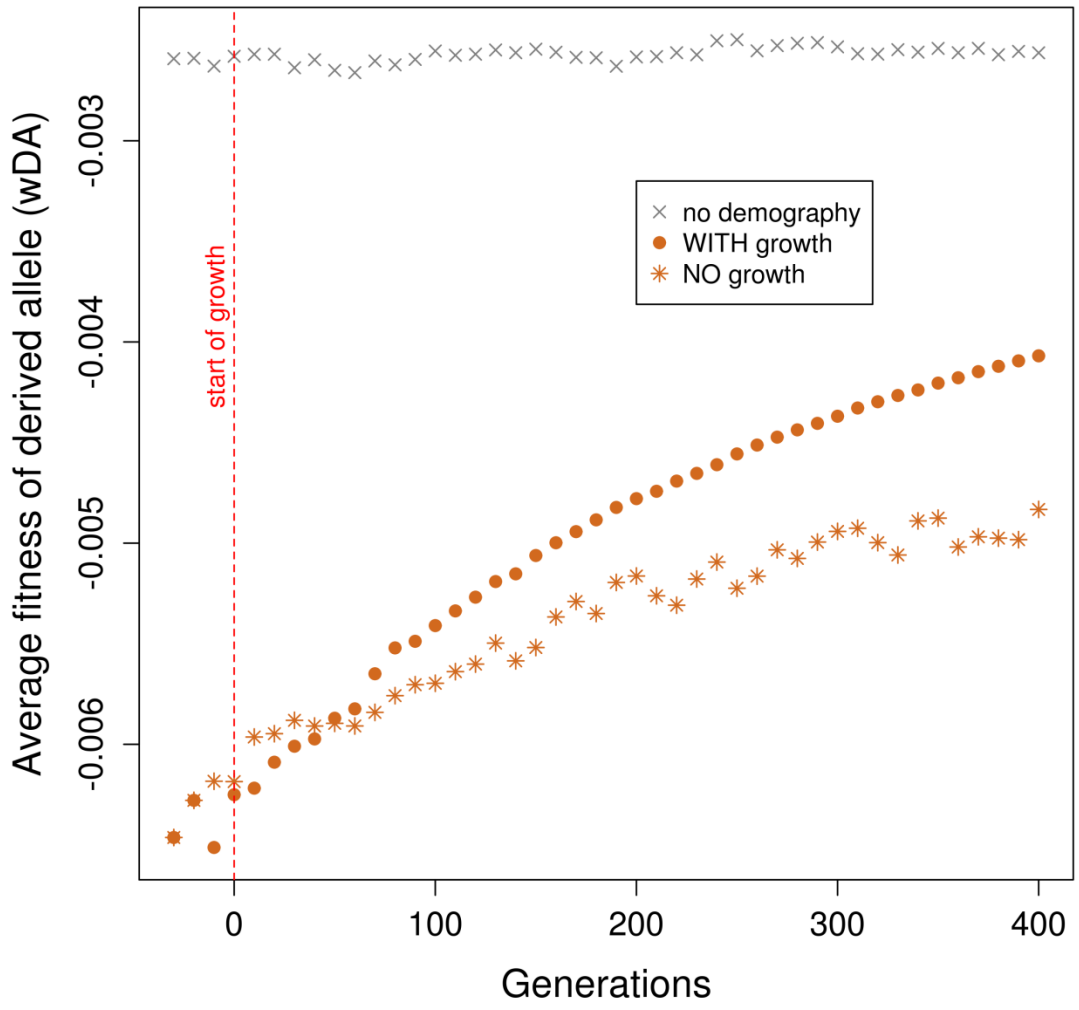


Figure S9 The mean fitness effect of the derived alleles in both demographic scenarios is lower than in a population without any demography. The growing and constant populations data are the same as presented in Figure 4. The population with no demography has a constant effective population size of 10,000 throughout history without any demographic events, thus differing from the constant size model by the absence of ancestral bottlenecks. The comparison of the constant size population and the population with no demography (gray) allows assessing the effect of past demographic events (bottlenecks). The average fitness of derived alleles in both populations with demographic events is lower than in the population with no demography. The increase in average fitness effect of the constant population size, although slower than the increase in the growing population, is explained by the progressive elimination of copies of derived alleles with low fitness effect that accumulated during the ancient bottlenecks. This effect is shown in empirical data in Lohmueller *et al.* (2008).

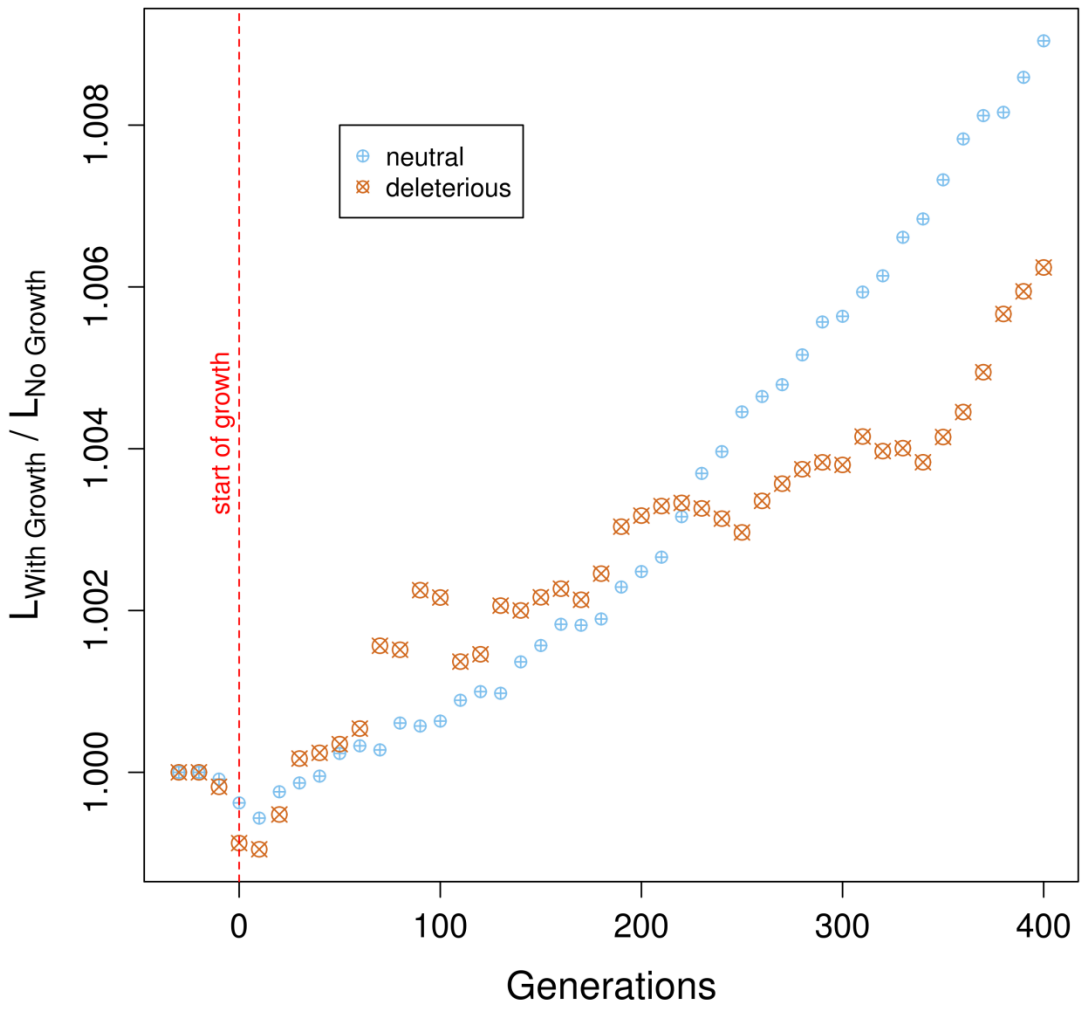


Figure S10 The number of mutations per chromosome is higher in a growing population. The number of mutations per chromosome (*L*) is defined as in Figure 5. Data shown is the ratio of *L* in a population with growth over *L* in a population without growth. Ratios above 1 indicate that *L* is higher under population expansion.

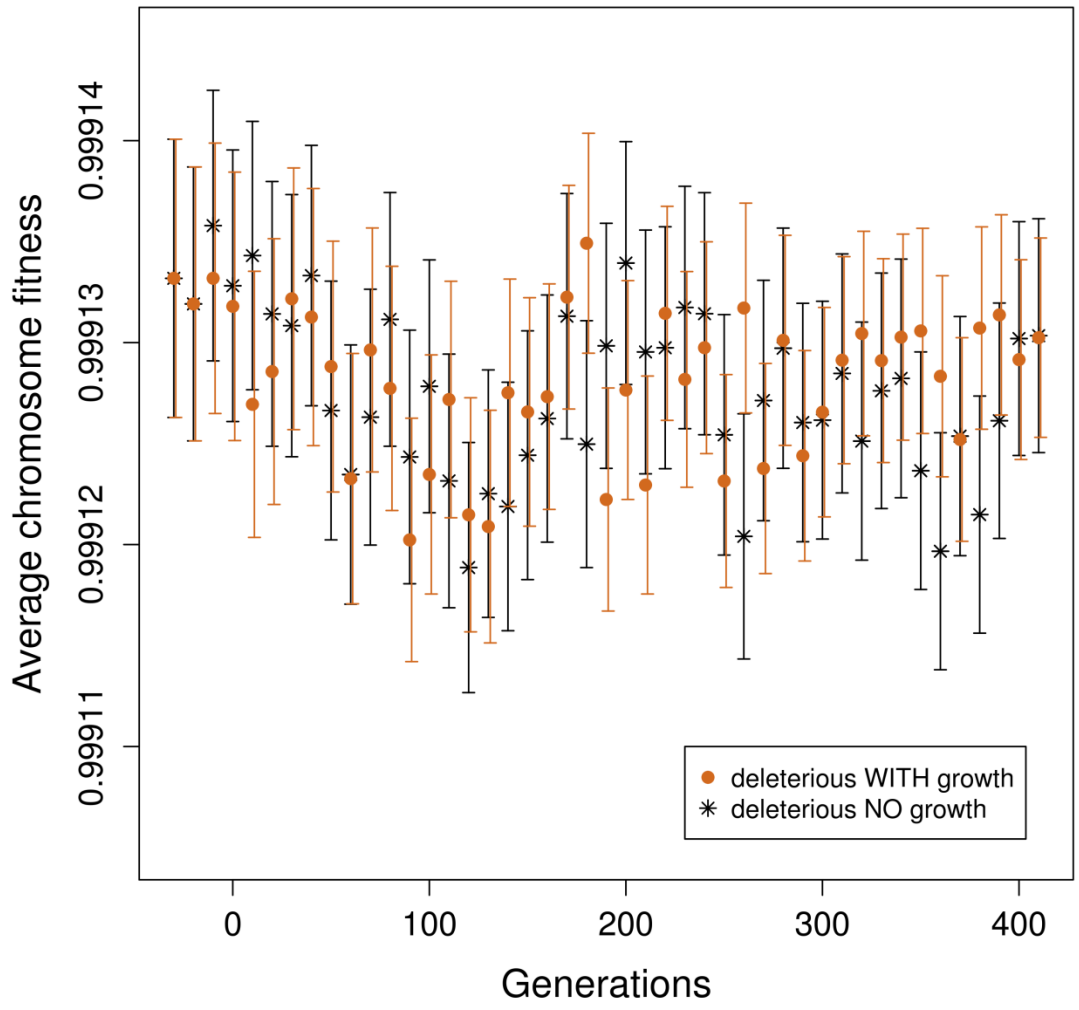


Figure S11 The average fitness of individuals is not different in the two demographic scenarios. The fitness of an individual chromosome relative to a fitness of 1 is defined as the product of selective coefficient of all mutations it carries. The average chromosome fitness, w_{CHR}, is defined as $\sum_{k=1}^{2Ne}\prod_{i=1}^{S}(1+s_{ik})\ /2Ne$, where s_{ik} is the selection coefficient of the derived allele at site i on chromosome *k*. In a growing population, individual chromosomes carry a larger number of mutations, but each one of them has on average a less deleterious effect, while in the population without growth an individual chromosome carries fewer mutations of larger fitness effect. Vertical bars represent the standard errors of the mean over 10,000 replicates.

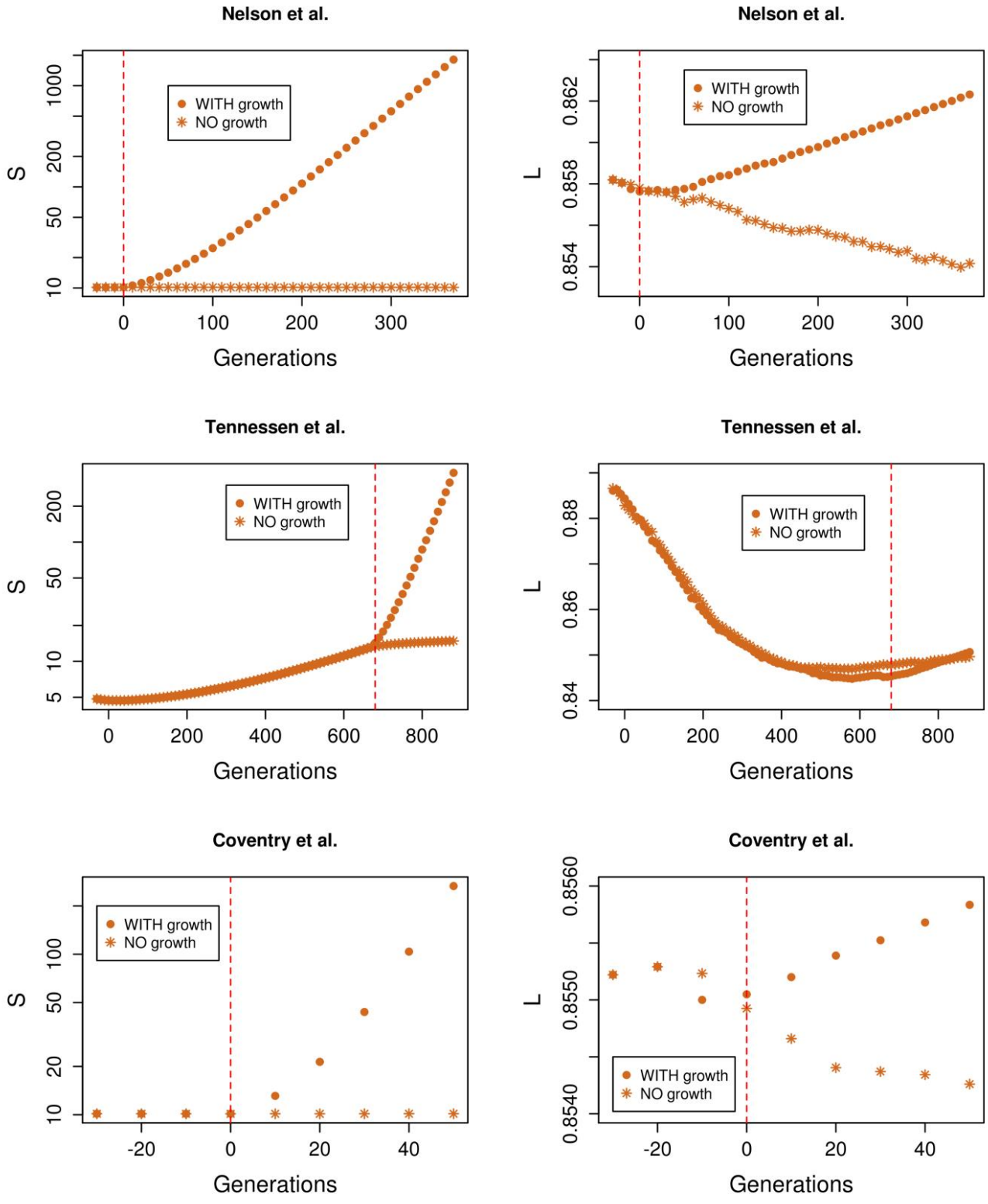


Figure S12 The number of segregating sites and the number of mutations per chromosome increase in growing populations assuming different models of European history. *S* and *L* are defined as in Figure 1 and Figure 5, respectively. Data are shown for a locus with deleterious mutations.

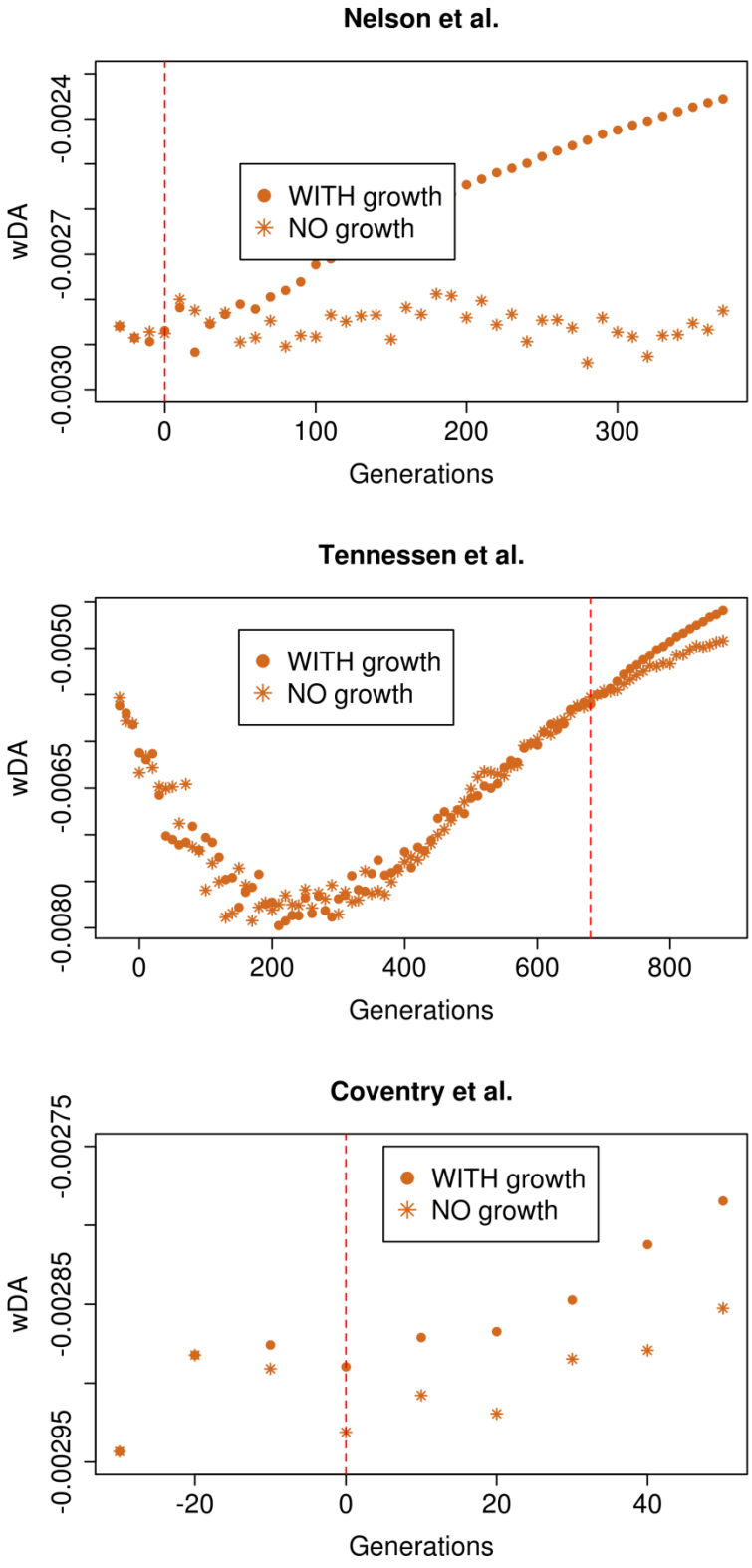


Figure S13 The average fitness effect of the alleles present in the population increases in growing populations assuming different models of European history. *wDA* is defined as in Figure 4.

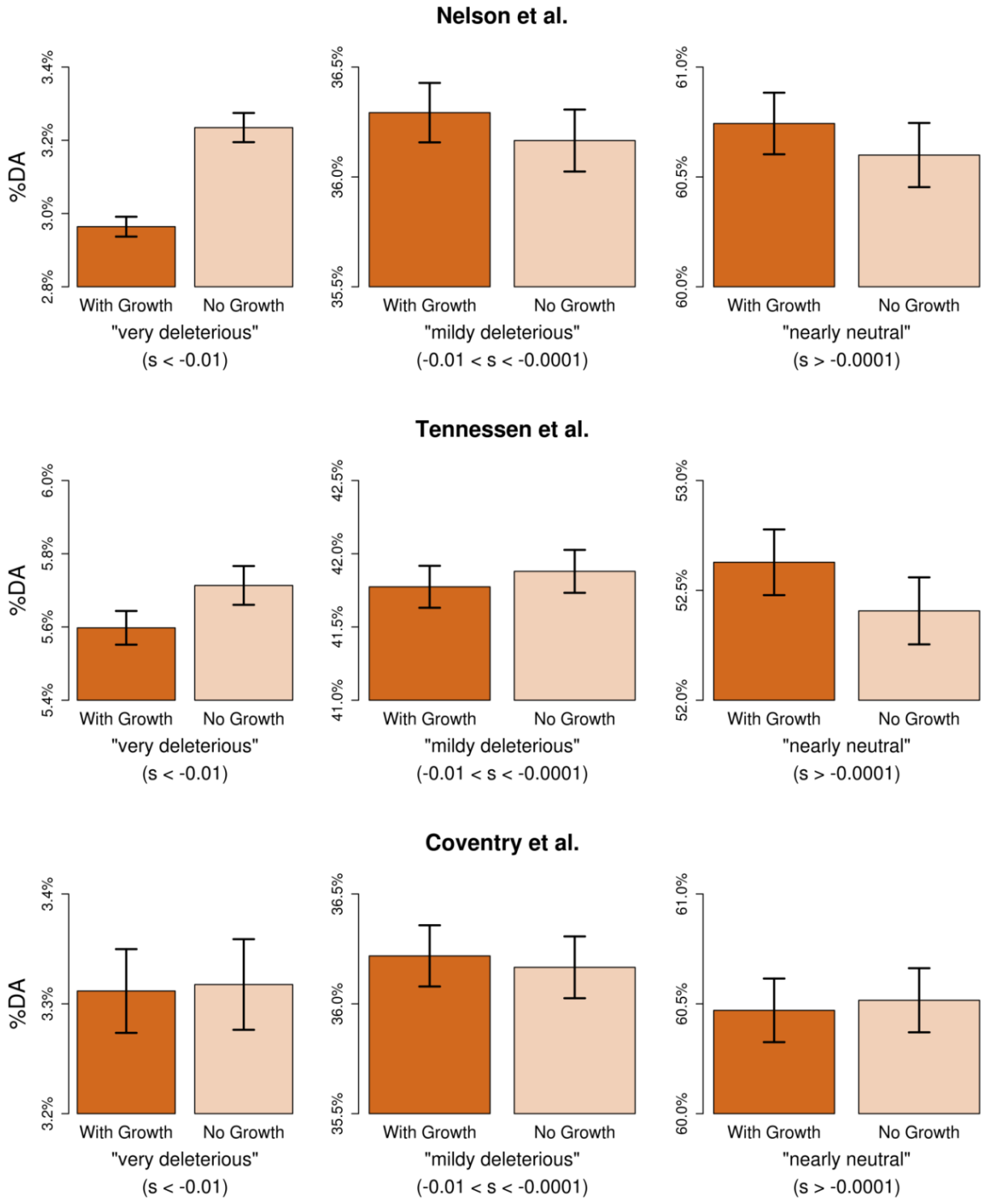


Figure S14 The percentage of sites in the most deleterious category of fitness decreases in growing populations assuming different models of European history. *%DA* and categories of fitness effect are defined as in Figure 3 and Figure S7. The growing population has a lower percentage of very deleterious mutations in all models but the one of Coventry *et al.*, where growth is extremely recent (56 generations) and its effect on very deleterious alleles is not significant.

File S1

Additional Results

Derived Allele Count (DAC) rather than allele frequency determines the probability of loss or transmission

In absence of selection, the probability for a segregating site of DAC = n to be lost or transmitted to the next generation follows a binomial process, where the probability of success (transmission) is $\frac{n}{2N_e}$ and the number of trials is $2N_e$ (where $2N_e$ is the number of chromosomes in the population), assuming the same number of individuals in the next generation.

To simplify, let us consider segregating sites with DAC = 1 (singletons). Each derived allele with DAC = 1 in a population of size N_e =10,000,000 has a frequency of only 1/10,000,000, which is much smaller than their frequency in a population of size N_e = 10,000. For each singleton taken individually, the probability of transmission to the next generation (probability of success in the binomial process) is lower in the larger population. However, the number of opportunities to be transmitted to the next generation (number of trials in the binomial process) increases, as more chromosomes are drawn from the current generation to participate to the next one. Consequently, the probability for one singleton to be lost is the same in the smaller population (N_e = 10,000) and in the larger population (N_e = 10,000,000). To see why this is so, consider that the probability of drawing zero copies from the binomial is Pr(X=0) = $[(N\text{-}1)/N]^N$. Using a normal approximation to the binomial distribution, we see that this value converges to 0.368 as N grows. This shows that for a segregating site, the probability of loss or transmission does not depend on the frequency of the derived allele but rather on its DAC. The same reasoning can be applied equally to other values of DAC. It also applies to a scenario of population growth, where the probability of losing a site still depends on its DAC rather than its frequency, and the number of trials is the population size at the next generation.

Effect of bottlenecks on the average fitness effect of a copy of derived allele

One additional population model was simulated (10,000 replicates) in order to test the effect of the bottlenecks in the two demographic scenarios studied here (*i.e.* the model of European history with growth and the model without growth). This additional population model has a constant effective size of 10,000 throughout history, with no demographic events at all. We refer to this population as the "no demography" population.

Both in the "no growth" and "with growth" models, the average selection coefficient of a copy of derived allele is much lower than in the model with no demography (Figure S9). This shows that the ancestral bottlenecks have reduced the average fitness of the

SNPs present in the population. Strong population contraction episodes can induce a drop on fitness through the random increase in frequency of deleterious mutations.

Command lines used to simulate the main demographic scenario

- Main model, with deleterious mutations. P0 and P1 split 420 generations ago, 20 generations before the beginning of the growth. During the last 400 generations, P1 grows to 10,000,000 and P0 remains at the constant size of 10,000.

./sfs_code 2 1 -N 1000 -t 0.0008 -n 1000 -L 1 5000 -a N -W 2 0 1 0.2 0.206 0.000365 -Td 0 P 0 0.0758 -Td 0.02 P 0 13.2 -Td 0.2 P 0 0.0769 -Td 0.207 P 0 13 -TS 0.215 0 1 -Tg 0.216 P 1 345.37501 -TE 0.236 --noSeq --trackTrajectory T 0.214 F output -o out

- Main model, with neutral mutations. P0 and P1 split 420 generations ago, 20 generations before the beginning of the growth. During the last 400 generations, P1 grows to 10,000,000 and P0 remains at the constant size of 10,000.

./sfs_code 2 1 -N 1000 -t 0.0008 -n 1000 -L 1 5000 -a N -W 0 -Td 0 P 0 0.0758 -Td 0.02 P 0 13.2 -Td 0.2 P 0 0.0769 -Td 0.207 P 0 13 -TS 0.215 0 1 -Tg 0.216 P 1 345.37501 -TE 0.236 --noSeq --trackTrajectory T 0.214 F output -o out

Additional demographic scenario

To test the robustness of our results to the demographic scenario assumed in the main text (main model), we repeated the analysis using 3 other published demographic scenarios involving a European model of ancient history and a final epoch of rapid exponential growth (Coventry et al., 2010; Nelson et al., 2012; Tennessen et al., 2012). Both the model of Coventry et al. (2010) and the model of Nelson et al. (2012) use the ancient demographic history described in Schaffner et al. (2005). The model of Tennessen et al. (2012) is based on the ancient demography of European history described in Gravel et al. (2011). The migration parameter of Schaffner et al. (2005) and Gravel et al. (2011) was ignored here, as done in the models by Tennessen et al. (2012), Nelson et al. (2012) and Coventry et al. (2010). We repeated the analysis applied to the main model, but only for a locus with deleterious mutations. The distribution of selective coefficients in the additional models was set to have the same mean (-0.28) as in the main model. Also identical to the main model, the populations with or without growth emerge from the split of a common ancient population to insure identical states when the comparison between growth and no growth starts (see main text). The populations with and without growth are followed for the whole duration of the growth, which varied according to the model. In all three additional models the final N_e is smaller

than in the main model, and therefore fewer mutations are outputted. To match the main model's level of information, we increased the number of replicates in the additional models to 50,000 replicates. The arguments used in SFS_code are the following:

- Nelson et al. (2012) model, with deleterious mutations:

./sfs_code 2 1 -N 1250 -t 0.0008 -n 1000 -L 1 5000 -a N -W 2 0 1 0.2 0.206 0.000292 -TN 0 2400 -TN 0.54 770 -Td 0.54004 0.07639 -Td 0.54404 13.1 -Td 0.6 0.3247 -Td 0.604 3.08 -TS 0.6648 0 1 -Tg 0.6652 P 1 422.441 -TE 0.68 --noSeq --trackTrajectory T 0.6636 F output -o out'

In this model, the final population sizes are N_e = 4,000,020 for the growing population and N_e = 7,700 for the population without growth.

- Tennessen et al. (2012) model, with deleterious mutations:

./sfs_code 2 1 -N 740 -t 0.0008 -n 1000 -L 1 5000 -a N -W 2 0 1 0.2 0.206 0.0004931 -TN 0 1447 -TN 0.2622 186 -TN 0.3378 104 -TS 0.3379 0 1 -Tg 0.3379 45.47 -Tg 0.3861 P 0 0 -Tg 0.3861 P 1 283.16 -TE 0.40 --noSeq --trackTrajectory T 0.3379 F output -o out

The model of Tennessen et al. (2012) uses two epochs of growth, a first one that is slow and a second one that is fast. The split occurs before the first growth epoch but the two populations grow at the same rate and at the end of the first epoch of growth, both populations have N_e = 9,210. Only one population undergoes the second (fast) epoch of growth and grows until N_e = 512,010, while the other one stops growing and remains at 9,210 individuals until the end of the simulation.

- Coventry et al. (2010) model, with deleterious mutations:

./sfs_code 2 1 -N 1250 -t 0.0008 -n 1000 -L 1 5000 -a N -W 2 0 1 0.2 0.206 0.000292 -TN 0 2400 -TN 0.54 770 -Td 0.54004 0.07639 -Td 0.54404 13.1 -Td 0.6 0.3247 -Td 0.604 3.08 -TS 0.6774 0 1 -Tg 0.678 P 1 2480.58 -TE 0.68 --noSeq --trackTrajectory T 0.6764 F output -o out

In this model, the final Ne of the population with growth is 1,100,020, while the population without growth remains as N_e = 7,700.

Time to the most recent common ancestor (t_{MRCA}) for a sample of two chromosome (n = 2)

- For a model without growth:

$$\Pr(0 < t < t_1) = \int_0^{t_1} \frac{1}{2N_a} e^{\frac{-t}{2N_a}} dt$$

$$\Pr(t_1 < t < t_2) = [1 - \Pr(0 < t < t_1)] \int_{t_1}^{t_2} \frac{1}{2N_a} e^{\frac{-(t-t_1)}{2N_a}} dt$$

$$\Pr(t_2 < t < t_3) = [1 - \Pr(0 < t < t_2)] \int_{t_2}^{t_3} \frac{1}{2N_{b_2}} e^{\frac{-(t-t_2)}{2N_{b_2}}} dt$$

$$\Pr(t_3 < t < t_4) = [1 - \Pr(0 < t < t_3)] \int_{t_3}^{t_4} \frac{1}{2N_a} e^{\frac{-(t-t_3)}{2N_a}} dt$$

$$\Pr(t_4 < t < t_5) = [1 - \Pr(0 < t < t_4)] \int_{t_4}^{t_5} \frac{1}{2N_{b_1}} e^{\frac{-(t-t_4)}{2N_{b_1}}} dt$$

$$E[t_{MRCA}(2)] = \int_0^{t_1} \frac{1}{2N_a} e^{\frac{-t}{2N_a}} t dt + [1 - \Pr(0 < t < t_1)] \int_{t_1}^{t_2} \frac{1}{2N_a} e^{\frac{-(t-t_1)}{2N_a}} t dt + [1 - \Pr(0 < t < t_2)] \int_{t_2}^{t_3} \frac{1}{2N_{b_2}} e^{\frac{-(t-t_2)}{2N_{b_2}}} t dt$$
$$+ [1 - \Pr(0 < t < t_3)] \int_{t_3}^{t_4} \frac{1}{2N_a} e^{\frac{-(t-t_3)}{2N_a}} t dt + [1 - \Pr(0 < t < t_4)] \int_{t_4}^{t_5} \frac{1}{2N_{b_1}} e^{\frac{-(t-t_4)}{2N_{b_1}}} t dt$$
$$+ [1 - \Pr(0 < t < t_5)] \int_{t_5}^{\infty} \frac{1}{2N_a} e^{\frac{-(t-t_5)}{2N_a}} t dt$$

- For a model with growth:

$$P(0 < t < t_1) = \int_0^{t_1} L(t) e^{-G(t)} dt$$

Where: $L(t) = \frac{1}{2N_a e^{g(t_1 - t)}}$ $\quad G(t) = \int_0^t \frac{1}{L(s)} ds$ g = rate of growth (*)

$$E[t_{MRCA}(2)] = \int_0^{t_1} L(t) e^{-G(t)} t dt + [1 - \Pr(0 < t < t_1)] \int_{t_1}^{t_2} \frac{1}{2N_a} e^{\frac{-(t-t_1)}{2N_a}} t dt + [1 - \Pr(0 < t < t_2)] \int_{t_2}^{t_3} \frac{1}{2N_{b_2}} e^{\frac{-(t-t_2)}{2N_{b_2}}} t dt$$
$$+ [1 - \Pr(0 < t < t_3)] \int_{t_3}^{t_4} \frac{1}{2N_a} e^{\frac{-(t-t_3)}{2N_a}} t dt + [1 - \Pr(0 < t < t_4)] \int_{t_4}^{t_5} \frac{1}{2N_{b_1}} e^{\frac{-(t-t_4)}{2N_{b_1}}} t dt$$
$$+ [1 - \Pr(0 < t < t_5)] \int_{t_5}^{\infty} \frac{1}{2N_a} e^{\frac{-(t-t_5)}{2N_a}} t dt$$

* From Griffiths and Tavaré (1994)

With the following parameters (time measures in generations):

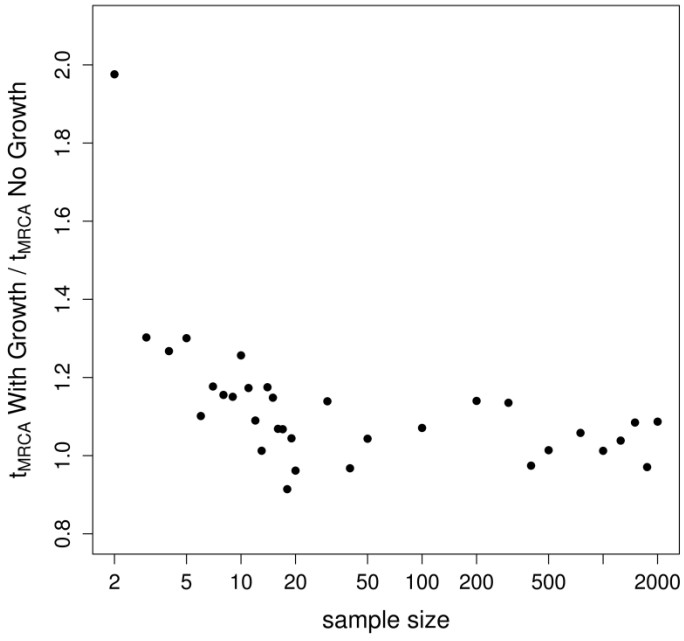
Parameter	Value
t_1	400
t_2	580
t_3	720
t_4	4320
t_5	4720
N_a	10,000
N_{b1}	757.57
N_{b2}	769.23
g	0.01726939

We obtain the following estimates for t_{MRCA} (**):

Model	Time (Generations)
E[noGrowth]	15264.76
E[Growth]	15526.00

** integrals numerically evaluated using: http://www.solvemymath.com/online_math_calculator/calculus/definite_integral/index.php

Using the two estimates calculated above, the difference in t_{MRCA} for a pair of chromosomes between the scenario with growth and without growth is 1.7%. This value is higher than the difference in the number of mutations per chromosomes (*L*) computed for the whole sample at neutral loci (Figure 5). This is because the difference in t_{MRCA} is maximal for a pair of samples and decreases with sample size. This effect is shown on the figure below: while the t_{MRCA} is expected to increase with sample size, the difference in t_{MRCA} decreases as sample size increases. The t_{MRCA} are computed over 1,000,000 replicates with the coalescent simulator ms (Hudson, 2002).



References

Coventry, A., Bull-Otterson, L.M., Liu, X., Clark, A.G., Maxwell, T.J., Crosby, J., Hixson, J.E., Rea, T.J., Muzny, D.M., Lewis, L.R., et al. (2010). Deep resequencing reveals excess rare recent variants consistent with explosive population growth. Nat Commun 1, 131.

Gravel, S., Henn, B.M., Gutenkunst, R.N., Indap, A.R., Marth, G.T., Clark, A.G., Yu, F., Gibbs, R.A., and Bustamante, C.D. (2011). Demographic history and rare allele sharing among human populations. Proc Natl Acad Sci U S A 108, 11983-11988.

Hudson, R.R. (2002). Generating samples under a Wright-Fisher neutral model of genetic variation. Bioinformatics 18, 337-338.

Nelson, M.R., Wegmann, D., Ehm, M.G., Kessner, D., St Jean, P., Verzilli, C., Shen, J., Tang, Z., Bacanu, S.A., Fraser, D., et al. (2012). An abundance of rare functional variants in 202 drug target genes sequenced in 14,002 people. Science 337, 100-104.

Schaffner, S.F., Foo, C., Gabriel, S., Reich, D., Daly, M.J., and Altshuler, D. (2005). Calibrating a coalescent simulation of human genome sequence variation. Genome Res 15, 1576-1583.

Tennessen, J.A., Bigham, A.W., O'Connor, T.D., Fu, W., Kenny, E.E., Gravel, S., McGee, S., Do, R., Liu, X., Jun, G., et al. (2012). Evolution and functional impact of rare coding variation from deep sequencing of human exomes. Science 337, 64-69.

Table S1 Characteristics of the mutations segregating in the population after 400 generations of growth or without growth.

	a)	Average number of copies per segregating site			b)	Average age of segregating sites (in generations)		
Bin of fitness		With Growth	No Growth	With Growth/No Growth		With Growth	No Growth	With Growth/No Growth
s<=-0.01		6.93	4.82	1.45		2.46	9.76	0.252
-0.01<s<=-0.0001		334.97	113.95	2.94		4.34	254.29	0.017
s>-0.0001		904.29	232.84	3.88		5.83	534.54	0.011

a) The number of copies at each segregating site is higher in the growing population for each category of fitness. However, the relative number of very deleterious variants compared to nearly neutral variants is much lower in the population with growth. For every copy of a very deleterious allele, there are almost 130 copies of a nearly neutral variant in the scenario with growth, while in the scenario with no growth there are 48 copies of a nearly neutral variant for each copy of a very deleterious variant.

b) In each category of fitness, the average age of the variants in the population with growth is more recent in all bins of fitness because the majority of the mutations segregating in the population have been introduced recently. Importantly, although the mutation rate is the same in both scenarios (meaning that the average fitness of the mutations introduced at each generation is the same in both scenarios), the average age of variants in the most deleterious category is younger in the model with growth because very deleterious mutations are eliminated more rapidly and more efficiently, as shown on figure S7.